

0590
1012

#7

OIEP

RAW SEQUENCE LISTING

DATE: 12/17/2001

PATENT APPLICATION: US/09/813,820

TIME: 10:19:16

Input Set : N:\Crf3\RULE60\09813820.txt

Output Set: N:\CRF3\12172001\I813820.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Hook, Magnus

7 Patti, Joseph M.

8 House-Pompeo, Karen

9 Sthanam, Narayana

10 Symersky, Jindrich

12 (ii) TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
13 AND METHODS OF USE

15 (iii) NUMBER OF SEQUENCES: 8

17 (iv) CORRESPONDENCE ADDRESS:

18 (A) ADDRESSEE: Arnold, White & Durkee

19 (B) STREET: P.O. Box 4433

20 (C) CITY: Houston

21 (D) STATE: Texas

22 (E) COUNTRY: U.S.

23 (F) ZIP: 77210

25 (v) COMPUTER READABLE FORM:

26 (A) MEDIUM TYPE: Floppy disk

27 (B) COMPUTER: IBM PC compatible

28 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

29 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

31 (vi) CURRENT APPLICATION DATA:

C--> 32 (A) APPLICATION NUMBER: US/09/813,820

C--> 33 (B) FILING DATE: 22-Mar-2001

34 (C) CLASSIFICATION:

36 (vii) PRIOR APPLICATION DATA:

37 (A) APPLICATION NUMBER: 08/856,253

38 (B) FILING DATE:

40 (viii) ATTORNEY/AGENT INFORMATION:

41 (A) NAME: Kitchell, Barbara S.

42 (B) REGISTRATION NUMBER: 33,928

43 (C) REFERENCE/DOCKET NUMBER: TAMK:193

45 (ix) TELECOMMUNICATION INFORMATION:

46 (A) TELEPHONE: (512) 418-3000

47 (B) TELEFAX: (512) 474-7577

50 (2) INFORMATION FOR SEQ ID NO: 1:

52 (i) SEQUENCE CHARACTERISTICS:

53 (A) LENGTH: 441 base pairs

54 (B) TYPE: nucleic acid

55 (C) STRANDEDNESS: single

56 (D) TOPOLOGY: linear

58 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

60 ATAACATCTG GGAATAAATC AACGAATGTT ACGGTTTCATA AAAGTGAAGC GGGAACAAGT 60

62 AGTGTTTTCT ATTATAAAAC GGGAGATATG CTACCAGAAG ATACGACACA TGTACGATGG 120

64 TTTTAAATA TTAACAATGA AAAAAGTTAT GTATCGAAAG ATATTACTAT AAAGGATCAG 180

66 ATTCAAGGTG GACAGCAGTT AGATTTAAGC ACATTAAACA TTAATGTGAC AGGTACACAT 240

ENTERED

RAW SEQUENCE LISTING

DATE: 12/17/2001

PATENT APPLICATION: US/09/813,820

TIME: 10:19:16

Input Set : N:\Crf3\RULE60\09813820.txt

Output Set: N:\CRF3\12172001\I813820.raw

```

68 AGCAATTATT ATAGTGGACA AAGTGCAATT ACTGATTTTG AAAAAGCCTT TCCAGGTTCT 300
70 AAAATAACTG TTGATAATAC GAAGAACACA ATTGATGTAA CAATTCCACA AGGCTATGGG 360
72 TCATATAATA GTTTTTCAAT TAACTACAAA ACCAAAATTA CGAATGAACA GCAAAAAGAG 420
74 TTTGTTAATA ATTCACAAGC T 441
77 (2) INFORMATION FOR SEQ ID NO: 2:
78 (i) SEQUENCE CHARACTERISTICS:
79 (A) LENGTH: 159 amino acids
80 (B) TYPE: amino acid
81 (C) STRANDEDNESS:
82 (D) TOPOLOGY: linear
83 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
84 Met Arg Gly Ser His His His His His His Gly Ser Ile Thr Ser Gly
85 1 5 10 15
86 Asn Lys Ser Thr Asn Val Thr Val His Lys Ser Glu Ala Gly Thr Ser
87 20 25 30
88 Ser Val Phe Tyr Tyr Lys Thr Gly Asp Met Leu Pro Glu Asp Thr Thr
89 35 40 45
90 His Val Arg Trp Phe Leu Asn Ile Asn Asn Glu Lys Ser Tyr Val Ser
91 50 55 60
92 Lys Asp Ile Thr Ile Lys Asp Gln Ile Gln Gly Gly Gln Gln Leu Asp
93 65 70 75 80
94 Leu Ser Thr Leu Asn Ile Asn Val Thr Gly Thr His Ser Asn Tyr Tyr
95 85 90 95
96 Ser Gly Gln Ser Ala Ile Thr Asp Phe Glu Lys Ala Phe Pro Gly Ser
97 100 105 110
98 Lys Ile Thr Val Asp Asn Thr Lys Asn Thr Ile Asp Val Thr Ile Pro
99 115 120 125
100 Gln Gly Tyr Gly Ser Tyr Asn Ser Phe Ser Ile Asn Tyr Lys Thr Lys
101 130 135 140
102 Ile Thr Asn Glu Gln Gln Lys Glu Phe Val Asn Asn Ser Gln Ala
103 145 150 155
104 (2) INFORMATION FOR SEQ ID NO: 3:
105 (i) SEQUENCE CHARACTERISTICS:
106 (A) LENGTH: 849 base pairs
107 (B) TYPE: nucleic acid
108 (C) STRANDEDNESS: single
109 (D) TOPOLOGY: linear
110 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
111 GACGATAAAA ATGGAAAAAT ACAAATGGT GACATGATTA AAGTGGCATG GCCGACAAGC 60
112 GGTACAGTAA AGATAGAGGG TTATAGTAAA ACAGTACCAT TAACTGTTAA AGGTGAACAG 120
113 GTGGGTCAAG CAGTTATTAC ACCAGACGGT GCAACAATTA CATTCAATGA TAAAGTAGAA 180
114 AAATTAAGTG ATGTTTCGGG ATTTGCAGAA TTTGAAGTAC AAGGAAGAAA TTTAACGCAA 240
115 ACAAATACTT CAGATGACAA AGTAGCTACG ATAACATCTG GGAATAAATC AACGAATGTT 300
116 ACGGTTTATA AAAGTGAAGC GGAACAAGT AGTGTTTTCT ATTATAAAAC GGGAGATATG 360
117 CTACCAGAAG ATACGACACA TGTACGATGG TTTTAAATA TTAACAATGA AAAAAGTTAT 420
118 GTATCGAAAG ATATTACTAT AAAGGATCAG ATCAAGGTG GACAGCAGTT AGATTTAAGC 480
119 ACATTAAACA TTAATGTGAC AGGTACACAT AGCAATTATT ATAGTGGACA AAGTGCAATT 540
120 ACTGATTTTG AAAAAGCCTT TCCAGGTTCT AAAATAACTG TTGATAATAC GAAGAACACA 600
121 ATTGATGTAA CAATTCCACA AGGCTATGGG TCATATAATA GTTTTTCAAT TAACTACAAA 660

```

RAW SEQUENCE LISTING

DATE: 12/17/2001

PATENT APPLICATION: US/09/813,820

TIME: 10:19:16

Input Set : N:\Crf3\RULE60\09813820.txt

Output Set: N:\CRF3\12172001\I813820.raw

```

150 ACCAAAATTA CGAATGAACA GCAAAAAGAG TTTGTTAATA ATTCACAAGC TTGGTATCAA      720
152 GAGCATGGTA AGGAAGAAGT GAACGGGAAA TCATTTAATC ATACTGTGCA CAATATTAAT      780
154 GCTAATGCCG GTATTGAAGG TACTGTAAAA GGTGAATTAA AAGTTTAAA ACAGGATAAA      840
156 GATACCAAG                                         849

```

159 (2) INFORMATION FOR SEQ ID NO: 4:

161 (i) SEQUENCE CHARACTERISTICS:

162 (A) LENGTH: 211 amino acids

163 (B) TYPE: amino acid

164 (C) STRANDEDNESS:

165 (D) TOPOLOGY: linear

167 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

169 Met Arg Gly Ser His His His His His His Gly Ser Asp Asp Lys Val
170 1 5 10 15
172 Ala Thr Ile Thr Ser Gly Asn Lys Ser Thr Asn Val Thr Val His Lys
173 20 25 30
175 Ser Glu Ala Gly Thr Ser Ser Val Phe Tyr Tyr Lys Thr Gly Asp Met
176 35 40 45
178 Leu Pro Glu Asp Thr Thr His Val Arg Trp Phe Leu Asn Ile Asn Asn
179 50 55 60
181 Glu Lys Ser Tyr Val Ser Lys Asp Ile Thr Ile Lys Asp Gln Ile Gln
182 65 70 75 80
184 Gly Gly Gln Gln Leu Asp Leu Ser Thr Leu Asn Ile Asn Val Thr Gly
185 85 90 95
187 Thr His Ser Asn Tyr Tyr Ser Gly Gln Ser Ala Ile Thr Asp Phe Glu
188 100 105 110
190 Lys Ala Phe Pro Gly Ser Lys Ile Thr Val Asp Asn Thr Lys Asn Thr
191 115 120 125
193 Ile Asp Val Thr Ile Pro Gln Gly Tyr Gly Ser Tyr Asn Ser Phe Ser
194 130 135 140
196 Ile Asn Tyr Lys Thr Lys Ile Thr Asn Glu Gln Gln Lys Glu Phe Val
197 145 150 155 160
199 Asn Asn Ser Gln Ala Trp Tyr Gln Glu His Gly Lys Glu Glu Val Asn
200 165 170 175
202 Gly Lys Ser Phe Asn His Thr Val His Asn Ile Asn Ala Asn Ala Gly
203 180 185 190
205 Ile Glu Gly Thr Val Lys Gly Glu Leu Lys Val Leu Lys Gln Asp Lys
206 195 200 205
208 Asp Thr Lys
209 210

```

212 (2) INFORMATION FOR SEQ ID NO: 5:

214 (i) SEQUENCE CHARACTERISTICS:

215 (A) LENGTH: 1500 base pairs

216 (B) TYPE: nucleic acid

217 (C) STRANDEDNESS: single

218 (D) TOPOLOGY: linear

220 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

222 GCACGAGATA TTTCAACAAC GAATGTTACA GATTAACTG TATCACCGTC TAAGATAGAA      60
224 GATGGTGGTA AAACGACAGT AAAAATGACG TTGCACGATA AAAATGGAAA AATACAAAAT      120
226 GGTGACATGA TTAAAGTGGC ATGGCCGACA AGCGGTACAG TAAAGATAGA GGGTTATAGT      180

```

RAW SEQUENCE LISTING

DATE: 12/17/2001

PATENT APPLICATION: US/09/813,820

TIME: 10:19:16

Input Set : N:\Crf3\RULE60\09813820.txt

Output Set: N:\CRF3\12172001\I813820.raw

```

228 AAAACAGTAC CATTAACTGT TAAAGGTGAA CAGGTGGGTC AAGCAGTTAT TACACCAGAC      240
230 GGTGCAACAA TTACATTCAA TGATAAAGTA GAAAAATTAA GTGATGTTTC GGGATTTGCA      300
232 GAATTTGAAG TACAAGGAAG AAATTTAACG CAAACAAATA CTTCAGATGA CAAAGTAGCT      360
234 ACGATAACAT CTGGGAATAA ATCAACGAAT GTTACGGTTC ATAAAAGTGA AGCGGGAACA      420
236 AGTAGTGTTC TCTATTATAA AACGGGAGAT ATGCTACCAG AAGATACGAC ACATGTACGA      480
238 TGGTTTTTAA ATATTAACAA TGAAAAAAGT TATGTATCGA AAGATATTAC TATAAAGGAT      540
240 CAGATTC AAG GTGGACAGCA GTTAGATTTA AGCACATTAA ACATTAATGT GACAGGTACA      600
242 CATAGCAATT ATTATAGTGG ACAAAGTGCA ATTACTGATT TTGAAAAAGC CTTTCCAGGT      660
244 TCTAAAAATAA CTGTTGATAA TACGAAGAAC ACAATTGATG TAACAATTCC ACAAGGCTAT      720
246 GGGTCATATA ATAGTTTTTC AATTAACTAC AAAACCAAAA TTACGAATGA ACAGCAAAAA      780
248 GAGTTTGTTA ATAATTCACA AGCTTGGTAT CAAGAGCATG GTAAGGAAGA AGTGAACGGG      840
250 AAATCATTTA ATCATACTGT GCACAATATT AATGCTAATG CCGGTATTGA AGGTACTGTA      900
252 AAAGGTGAAT TAAAAGTTTT AAAACAGGAT AAAGATACCA AGGCTCCTAT AGCTAATGTA      960
254 AAATTTAAAC TTTCTAAAAA AGATGGATCA GTTGTAAGG ACAATCAAAA AGAAATTGAG     1020
256 ATTATAACAG ATGCAAACGG TATTGCTAAT ATTAAAGCGT TGCCTAGTGG AGACTATATT     1080
258 TTAAGAGAAA TAGAGGCGCC ACGACCGTAT ACATTTGATA AGGATAAAGA ATATCCGTTT     1140
260 ACTATGAAAG ATACAGATAA TCAGGGATAT TTTACGACTA TTGAAAATGC AAAAGCGATA     1200
262 GAAAAAACAA AAGATGTTTC TGCTCAAAAAG GTTTGGGAAG GCACTCAAAA AGTGAAACCA     1260
264 ACGATTTATT TCAAGTTGTA CAAACAAGAT GACAATCAAAA ATACAACACC AGTAGACAAA     1320
266 GCAGAGATTA AAAAATTAGA AGATGGAACG ACAAAGTGTA CATGGTCTAA TCTTCCGGAA     1380
268 AATGACAAAA ATGGCAAGGC TATTAAATAT TTAGTTAAAG AAGTAAATGC TCAAGGTGAA     1440
270 GATACAACAC CAGAAGGATA TACTAAAAAA GAAATGGTT TAGTGGTTAC TAATACTGAA     1500

```

273 (2) INFORMATION FOR SEQ ID NO: 6:

275 (i) SEQUENCE CHARACTERISTICS:

276 (A) LENGTH: 512 amino acids

277 (B) TYPE: amino acid

278 (C) STRANDEDNESS:

279 (D) TOPOLOGY: linear

281 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```

283 Met Arg Gly Ser His His His His His His Gly Ser Ala Arg Asp Ile
284 1 5 10 15
286 Ser Ser Thr Asn Val Thr Asp Leu Thr Val Ser Pro Ser Lys Ile Glu
287 20 25 30
289 Asp Gly Gly Lys Thr Thr Val Lys Met Thr Phe Asp Asp Lys Asn Gly
290 35 40 45
292 Lys Ile Gln Asn Gly Asp Met Ile Lys Val Ala Trp Pro Thr Ser Gly
293 50 55 60
295 Thr Val Lys Ile Glu Gly Tyr Ser Lys Thr Val Pro Leu Thr Val Lys
296 65 70 75 80
298 Gly Glu Gln Val Gly Gln Ala Val Ile Thr Pro Asp Gly Ala Thr Ile
299 85 90 95
301 Thr Phe Asn Asp Lys Val Glu Lys Leu Ser Asp Val Ser Gly Phe Ala
302 100 105 110
304 Glu Phe Glu Val Gln Gly Arg Asn Leu Thr Gln Thr Asn Thr Ser Asp
305 115 120 125
307 Asp Lys Val Ala Thr Ile Thr Ser Gly Asn Lys Ser Thr Asn Val Thr
308 130 135 140
310 Val His Lys Ser Glu Ala Gly Thr Ser Ser Val Phe Tyr Tyr Lys Thr
311 145 150 155 160

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/813,820

DATE: 12/17/2001

TIME: 10:19:16

Input Set : N:\Crf3\RULE60\09813820.txt

Output Set: N:\CRF3\12172001\I813820.raw

```

313   Gly Asp Met Leu Pro Glu Asp Thr Thr His Val Arg Trp Phe Leu Asn
314                               165                               170                               175
316   Ile Asn Asn Glu Lys Ser Tyr Val Ser Lys Asp Ile Thr Ile Lys Asp
317                               180                               185                               190
319   Gln Ile Gln Gly Gly Gln Gln Leu Asp Leu Ser Thr Leu Asn Ile Asn
320                               195                               200                               205
322   Val Thr Gly Thr His Ser Asn Tyr Tyr Ser Gly Gln Ser Ala Ile Thr
323                               210                               215                               220
325   Asp Phe Glu Lys Ala Phe Pro Gly Ser Lys Ile Thr Val Asp Asn Thr
326   225                               230                               235                               240
328   Lys Asn Thr Ile Asp Val Thr Ile Pro Gln Gly Tyr Gly Ser Tyr Asn
329                               245                               250                               255
331   Ser Phe Ser Ile Asn Tyr Lys Thr Lys Ile Thr Asn Glu Gln Gln Lys
332                               260                               265                               270
334   Glu Phe Val Asn Asn Ser Gln Ala Trp Tyr Gln Glu His Gly Lys Glu
335                               275                               280                               285
337   Glu Val Asn Gly Lys Ser Phe Asn His Thr Val His Asn Ile Asn Ala
338   290                               295                               300
340   Asn Ala Gly Ile Glu Gly Thr Val Lys Gly Glu Leu Lys Val Leu Lys
341   305                               310                               315                               320
343   Gln Asp Lys Asp Thr Lys Ala Pro Ile Ala Asn Val Lys Phe Lys Leu
344                               325                               330                               335
346   Ser Lys Lys Asp Gly Ser Val Val Lys Asp Asn Gln Lys Glu Ile Glu
347   340                               345                               350
349   Ile Ile Thr Asp Ala Asn Gly Ile Ala Asn Ile Lys Ala Leu Pro Ser
350   355                               360                               365
352   Gly Asp Tyr Ile Leu Lys Glu Ile Glu Ala Pro Arg Pro Tyr Thr Phe
353   370                               375                               380
355   Asp Lys Asp Lys Glu Tyr Pro Phe Thr Met Lys Asp Thr Asp Asn Gln
356   385                               390                               395                               400
358   Gly Tyr Phe Thr Thr Ile Glu Asn Ala Lys Ala Ile Glu Lys Thr Lys
359                               405                               410                               415
361   Asp Val Ser Ala Gln Lys Val Trp Glu Gly Thr Gln Lys Val Lys Pro
362   420                               425                               430
364   Thr Ile Tyr Phe Lys Leu Tyr Lys Gln Asp Asp Asn Gln Asn Thr Thr
365   435                               440                               445
367   Pro Val Asp Lys Ala Glu Ile Lys Lys Leu Glu Asp Gly Thr Thr Lys
368   450                               455                               460
370   Val Thr Trp Ser Asn Leu Pro Glu Asn Asp Lys Asn Gly Lys Ala Ile
371   465                               470                               475                               480
373   Lys Tyr Leu Val Lys Glu Val Asn Ala Gln Gly Glu Asp Thr Thr Pro
374   485                               490                               495
376   Glu Gly Tyr Thr Lys Lys Glu Asn Gly Leu Val Val Thr Asn Thr Glu
377   500                               505                               510

```

380 (2) INFORMATION FOR SEQ ID NO: 7:

382 (i) SEQUENCE CHARACTERISTICS:

383 (A) LENGTH: 345 amino acids

384 (B) TYPE: amino acid

385 (C) STRANDEDNESS:

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/813,820

DATE: 12/17/2001

TIME: 10:19:17

Input Set : N:\Crf3\RULE60\09813820.txt

Output Set: N:\CRF3\12172001\I813820.raw

L:32 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:33 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]